# EXHIBIT 42

Cell, Vol. 16, 787-796, April 1979, Copyright © 1979 by MIT

# A Family of Short, Interspersed Repeat Sequences at the 5′ End of a Set of Dictyostelium Single-Copy mRNAs

Alan R. Kimmel and Richard A. Firtel
Department of Biology, B-022
University of California, San Diego
La Jolla, California 92093

## Summary

**M4 is a recombinant plasmid containing Dictyostelium nuclear DNA. The restriction map has been determined (K. Kindle and R. Firtel, manuscript submitted) and the four major fragments have been subcloned in pBR322. M4 is >90% single-copy; however, it does contain a short (300 bp) sequence repeated ~100 times in the Dictyostelium genome. The short repeat sequence is interspersed between single-copy regions. 1-1.5% of total vegetative poly(A)$^+$ mRNA hybridizes to the repeat, but only 10% of the hybridization is resistant to low levels of RNAase. The mRNA is heterogeneous in size and 90% of the mass is complementary to Dictyostelium single-copy DNA. We have also shown that one of the single-copy regions adjacent to the repeat is complementary to a low abundance class mRNA (0.01% of total mRNA). Sandwich hybridization experiments show that unlabeled RNA that hybridizes to this single-copy region will also hybridize $^{32}$P repeat. Exonuclease studies indicate that the repeat and single-copy regions are part of a single transcription unit which encodes a 1.2 kb mRNA. The 5′ end of this mRNA is complementary to the short repeat sequence and the remainder of the mRNA is transcribed from single-copy DNA. Furthermore, RNA excess hybridization to $^{32}$P-labeled separated strands of the repeat suggests that the repeat sequence is transcribed asymmetrically into poly(A)$^+$ RNA. We have proposed a model in which this repeat sequence is adjacent to ~100 different single-copy genes and is transcribed to produce different mRNA molecules carrying the common repeat sequence at the 5′ end.**

## Introduction

Approximately 50% of the Dictyostelium genome consists of short repeat sequences interspersed with single-copy DNA (Firtel and Kindle, 1975) in an organization similar to that of most metazoans (Davidson et al., 1975a). Linkage hybridization studies indicate that ~60% of hnRNA and ~25% of mRNA molecules in Dictyostelium are transcribed from both single-copy and r peat DNA (Firtel and Lodish, 1973; Firtel, Kindle and Huxley, 1976b). Moreover, other data suggest that the 5′ ends of both hnRNA and mRNA are enriched in transcribed repeat sequences (Firtel and L dish, 1973). Such studies using total DNA and mixed populations of mRNA are difficult to interpret. Clearly, the observed sequence organization of DNA and RNA must be related to the organization and transcription of individual genes. We have begun to study the organization of short repeat sequences by molecular cloning.

M4 is a recombinant plasmid containing a 4.5 kb nuclear Dictyostelium DNA fragment inserted into the Eco RI site of pMB9 by poly(dA-dT) homopolymer extension (K. Kindle and R. Firtel, manuscript submitted). M4 was originally selected for hybridization to a large fraction of $^{32}$P-mRNA by the Grunstein and Hogness (1975) colony filter hybridization technique. 1-1.5% of total $^{32}$P-vegetative mRNA hybridizes to M4, and 90% of the hybridization is sensitive to a low level of RNAase. The mRNA complementary to M4 is heterogeneous in size. DNA excess hybridization kinetics indicate that ~90% of the M4 insert is complementary to single-copy DNA although a repetitive fraction is clearly present. This conclusion is further supported by analyzing the hybridization of an M4 probe to restriction endonuclease digests of Dictyostelium DNA on DNA blot filters (Southern, 1975). It was suggested that the repeat sequence is transcribed onto a population of mRNAs (K. Kindle and R. Firtel, manuscript submitted).

We present evidence that M4 contains a transcription unit composed of a short repeat sequence plus adjacent single-copy DNA. The repeat sequence is transcribed onto the 5′ end of the mRNA; the remainder of the mRNA is complementary to single-copy DNA. Other data suggest that a large fraction of the members of that repeat sequence family are transcribed along wih their adjacent single-copy regions. Our results indicate that all members of the repeat sequence are probably transcribed from the same strand. This asymmetric transcription is discussed in relation to the possible function of such short repeat sequences in Dictyostelium.

## Results

### Construction of Subclones from M4 Restriction Fragments

The restriction map of plasmid M4 DNA is shown in Figure 1. Each band was individually subcloned to obtain large amounts of individual restriction fragments free from contamination with other sequences. Bands labeled 1, 2 and 3 were purified and inserted into the Pst 1 site within the amp$^R$ gene of pBR322. The isolated M4 fragments were extended with poly(dC) and combined with pBR322 which had been digested with Pst I and extended with poly(dG). The tailing of Pst I restriction sites with poly(dG) reconstructs the Pst I sites; thus fragments inserted into a cloning vehicle by this method can be excised free of vehicle sequences using Pst I (A. Otsuka, manuscript submitted; W. Rowekamp and R. Firtel, manuscript in

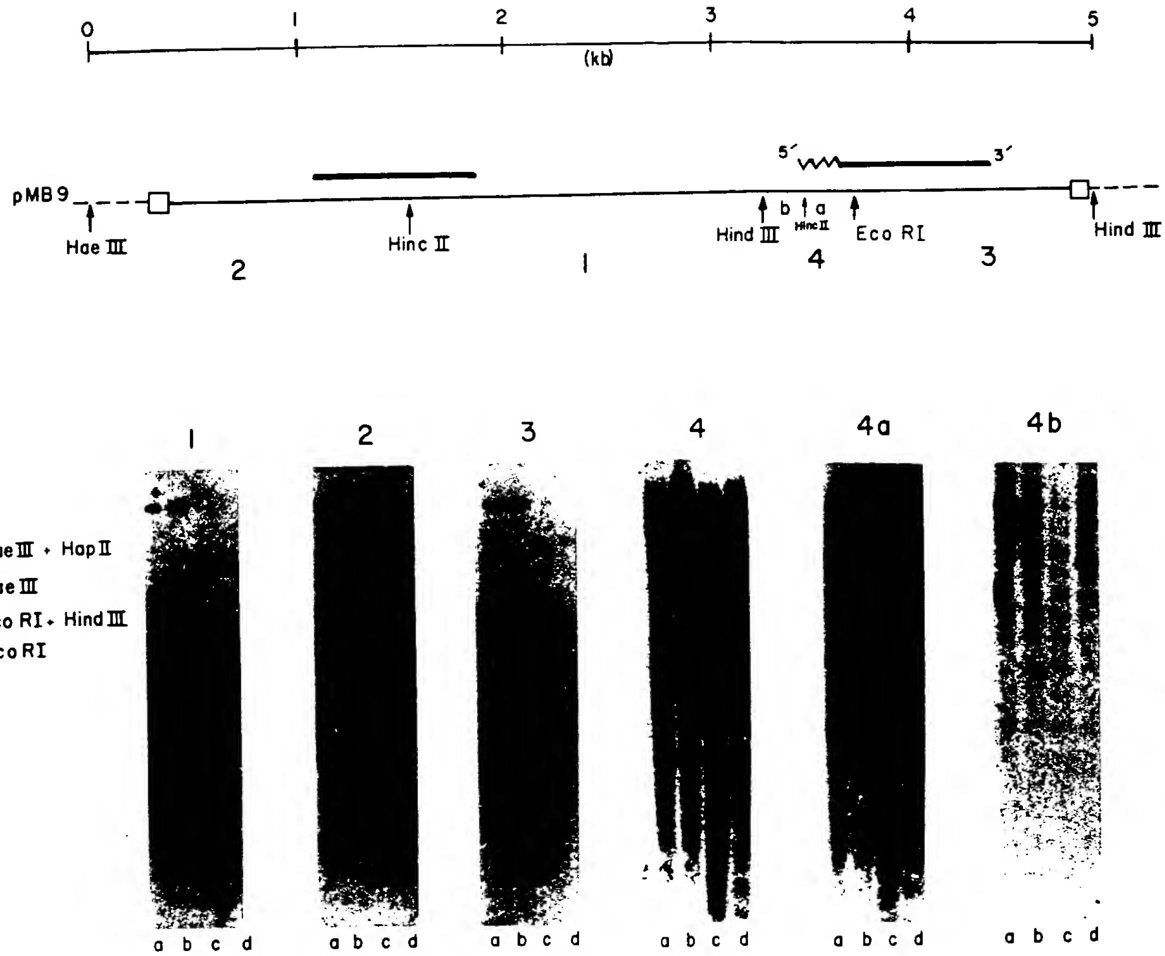


Figure 1. Location of Single-Copy and Repeat Sequences on Restriction Map of M4

M4 contains a 4.5 kb fragment of Dictyostelium DNA inserted into the Eco RI site of pMB9 by poly(dA-dT) homopolymer extension as described previously (K. Kindle and K. Firtel, manuscript submitted). There is a deletion in the pMB9 vehicle from the Eco RI site to near the Hind III site. The bars above the restriction map indicate the regions transcribed into mRNA (this paper). The polarity of the band 4-3 mRNA is indicated.

Bands 1-4 were purified from their respective subclones. Fragments 4a and 4b were prepared from subclone 4 using appropriate restriction digestions (see restriction maps). Probes were labeled in vitro by nick translation with $\alpha$-$^{32}$P-dXTPs and hybridized to various restriction digests of Dictyostelium DNA which had been size-fractionated on agarose gels and transferred to nitrocellulose filters. The filters were autoradiographed after appropriate washes. Details are discussed in Experimental Procedures.

preparation). M4 band 4 was subcloned into the Eco RI and Hind III site of pBR322 using T4 DNA ligase.

## M4 Contains a Short Interspersed Repeat Sequence

K. Kindle and R. Firtel (manuscript submitted) have shown that a small fraction (~10%) of the M4 plasmid DNA insert is repeated in the Dictyostelium genome. To locate the position of the repeat within M4, in vitro labeled $^{32}$P probes of isolated subcloned bands 1-4 were hybridized to DNA blot filters of Dictyostelium nuclear DNA digested with various restriction endonucleases and size-fractionated on agarose gels (Southern, 1975). As shown in Figure 1, bands 1, 2 and 3 show strong hybridization to only one restriction fragment in each of the digests. It is therefore suggested that each band is present in a single copy per genome. DNA excess hybridization kinetics of Dictyostelium nuclear DNA with nick-translated hybridization probes of bands 1, 2 and 3 confirm this conclusion (data not shown). It should also be noted that these hybridizations are entirely consistent with the M4 restriction map. Neither Hae III nor Hap II cleave the M4 insert, and as expected, bands 1, 2 and 3 are all located on the same Hae III genome fragment. Eco RI and Hind III sites, however, separate band 3 from bands 1 and 2 on M4. In the genome, band 3 hybridizes to a genome fragment of different size (see slots c and d) than do bands 1 and 2. Closer examination of the hybridization of band 3 to two of the Dictyostelium digests shown in slots c and d indicates minor hybridization to a second band. This hybridization has

been mapped to the segment of the Dictyostelium rDNA (Eco RI band III; Cockburn, Newkirk and Firtel, 1976a) which contains the coding region for the 5S ribosomal RNA and the nontranscribed spacer region extending to the 3′ end of the 25S ribosomal RNA coding region (Maizels, 1976; Cockburn et al., 1976a, 1976b; Cockburn, Taylor and Firtel, 1978). This minor hybridization is not seen in slots a and b since the Hae III and Hap II fragments of this region are too small to be detected in this gel system. We estimate that hybridization to the 200 copies of this region of the ribosomal DNA repeat occurs approximately 1000 times less often per sequence copy than does hybridization to the single-copy sequence. The significance of this minor hybridization is not yet understood.

In contrast to the pattern observed with bands 1–3, the 0.45 kb band 4 hybridizes to more than 50 restriction fragments of Dictyostelium DNA. Band 4 was cleaved with Hinc II, and the two fragments 4a (0.26 kb) and 4b (0.19 kb) were isolated, labeled by nick translation (Schachat and Hogness, 1973; Maniatis, Jeffrey and Kleid, 1975) and hybridized to genomic DNA Southern blots. It can be seen that fragment 4a shows a repeated sequence hybridization pattern while 4b is mostly single-copy, although the repeat pattern of hybridization is discernible upon longer exposure. The M4 0.45 kb Hind III-Eco RI genome fragment in slot c is not detected in this gel system. These results demonstrate that Hinc II cleaves within the repeat sequence; both 4a and 4b contain a repeat sequence although the majority of the 4b sequence is probably single-copy. To quantitate the repetition frequency of band 4, nick-translated band 4 was hybridized to a vast excess of Dictyostelium genomic DNA. Figure 2 shows that ~60% of band 4 hybridizes at a $Cot_{1/2}$ of ~1, while single-copy DNA hybridizes at a $Cot_{1/2}$ of ~120 (Firtel and Bonner, 1972). These data show that band 4 is repeated ~100 times in the genome. It should be noted that no single-copy sequence hybridization is observed with this probe. Single-copy sequences on labeled fragments which also contain repeat sequences may not be detected by our hybridization procedure. At relatively low Cot values (that is, <10), the single-copy regions would exist as tails on hybrid molecules. Smith, Britten and Davidson (1975) have shown that the nucleation rate for single-stranded regions on fragments which also contain duplexes is significantly retarded. In addition, some nick-translated fragments composed entirely of single-copy sequences (which are at the ends of band 4 and comprise less than one fourth of its total length) may be too short to form assayable hybrids or may show a retardation in the rate of hybridization as a result of their shorter length. We conclude that M4 DNA contains a short (~300 bp) repeat sequence interspersed with single-copy DNA in the genome.

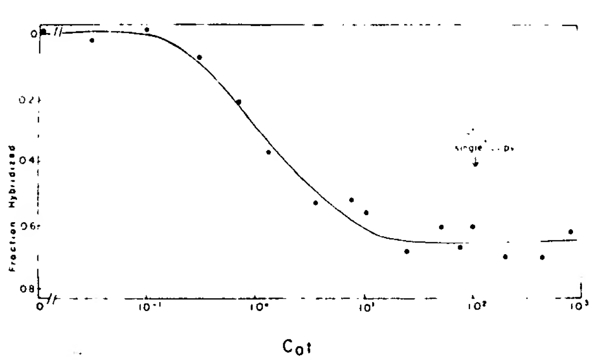


Figure 2. Hybridization Kinetics of $^{32}P$ Band 4 with Vast Excess of Dictyostelium Nuclear DNA

Nick-translated $^{32}P$ band 4 (spec. act. ~$3 \times 10^7$ cpm/μg) was hybridized in solution with a 100 fold single-copy sequence excess of nuclear DNA. Hybridizaton was assayed by S1 digestion and each point was corrected for S1 cleavage of unhybridized tails (Morrow, 1974). $^3H$ nuclear DNA tracer was included in each reaction and the $Cot_{1/2}$ of single-copy DNA is indicated. No renaturation of $^{32}P$ band 4 was detected in control hybridizations without Dictyostelium driver DNA (see Kindle and Firtel, 1978). Values were expressed as equivalent Cot (Britten, Graham and Neufeld, 1974).

## Band 4 Repeat Sequence Hybridizes ~1% of Total Poly(A)$^+$ mRNA

To determine which of the M4 fragments are complementary to mRNA, DNA from each M4 subclone was immobilized on nitrocellulose filters and hybridized with $^{32}P$-pulse-labeled poly(A)$^+$ mRNA (Kindle and Firtel, 1978). As shown in Table 1, ~1.2% of vegetative pulse-labeled mRNA hybridizes to total M4 plasmid DNA. When the filters are treated with RNAase,

Table 1. Percentages of Vegetative $^{32}P$-mRNA Which Hybridize to M4 Clones

| Clone | −RNAase | +RNAase |
|---|---|---|
| Total M4 | 1.2 | 0.2 |
| Band 1 | 0.1 | 0.05 |
| Band 2 | 0.1 | 0.07 |
| Band 3 | 0.02 | 0.01 |
| Band 4 | 1.0 | 0.1 |

DNA from each M4 clone was purified and immobilized on nitrocellulose filters. The filters were hybridized with in vivo labeled $^{32}P$-mRNA (Kindle and Firtel, 1978). Control filters containing only pMB9 or pBR322 vehicle DNA were included in each reaction to monitor nonspecific binding. Filters were washed extensively at criterion until no change in nonspecific binding was observed as measured by Cerenkov radiation. Filters were counted and then treated with a low level of RNAase A (0.05 μg/ml) for 1 hr at room temperature, washed and recounted. Values are expressed as the percentage of input radioactivity that hybridizes to specific cloned DNA, and are not corrected for nonspecific binding. The results are the averages of three independent experiments with ~15% variation. Similar values are obtained using whole cell $^{32}P$-poly(A)$^+$ RNA. Hybridizations were in DNA excess; 3 μg DNA per filter. Input Cerenkov radiation: total M4, band 4 = ~$0.5 \times 10^6$ cpm; band 1, band 2 = ~$5 \times 10^6$ cpm; band 3 = ~$20 \times 10^6$ cpm. Nonspecific pBR322 binding was ~200 cpm per filter, except in band 3 hybridization, which was 500–800 cpm per filter.

80–90% of the hybridized RNA is removed, as previously shown by K. Kindle and R. Firtel (manuscript submitted). Subclones 1 and 2 both hybridize ~0.1% of the mRNA and approximately 50% of the hybridization is resistant to RNAase. Only ~0.02% of total mRNA hybridizes to subcloned band 3. This level of hybridization is reproducible and is consistently greater than the control hybridization to filters carrying only pBR322 vehicle DNA. To quantitate the level of hybridization to band 3 more accurately, nick-translated band 3 DNA was hybridized with a vast excess of vegetative poly(A)$^+$ RNA. It can be seen in Figure 3 that band 3 hybridizes with a $Rot_{1/2}$ of 20. Assuming a $Rot_{1/2}$ of $\sim 1 \times 10^{-3}$ for a pure 1 kb mRNA (see legend to Figure 3), the RNA driving the hybridization represents 0.005% of the vegetative poly(A)$^+$ RNA, a value comparable with the data presented in Table 1.

In contrast to the relatively low levels of hybridization of mRNA to the single-copy subcloned fragments 1–3, band 4 (containing the repeat sequence) hybridizes ~1% of mRNA, only 10% of which is resistant to low levels of RNAase. These data suggest that only a small fraction (10–20%) of the actual sequences in the hybridizing RNA molecules are complementary to band 4 (or to total M4) DNA and that 80–90% of the sequences are unhybridized RNAase-sensitive tails. Southern blots of subclone 4 cleaved with Hinc II and hybridized with in vivo $^{32}$P-mRNA show that both 4a and 4b contain sequences complementary to a large fraction of vegetative mRNA. We conclude that a large fraction of vegetative mRNA hybridizes to the short interspersed repeat sequence in M4.

## Repeat Sequence Hybridizes a Heterogeneous Size Population of mRNA

The results shown in Table 1 indicate that the four subcloned restriction fragments isolated from M4 DNA are complementary to different quantities of mRNA. To examine the size distribution of the complementary RNA, $^{32}$P-mRNA was hybridized to total M4 plasmid DNA and plasmid DNA from the four subclones; the hybridized mRNA was eluted and electrophoresed on urea-SDS-polyacrylamide gels (Spradling, Pardue and Penman, 1977), and the RNA was identified by autoradiography. As shown in Figure 4, RNA complementary to total M4 migrates as a heterogeneous population although a specific mRNA of ~0.9 kb can be observed. Similar results have been obtained by K. Kindle and R. Firtel (manuscript submitted). Subcloned bands 1 and 2 hybridize only the 0.9 kb mRNA.

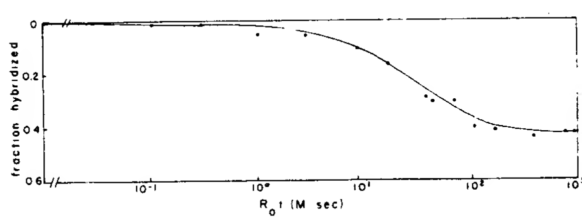


Figure 3. Hybridization Kinetics of $^{32}$P Band 3 with Vast Excess of Vegetative Poly(A)$^+$ RNA

$^{32}$P band 3 (spec. act. $\sim 4 \times 10^7$ cpm/µg) was hybridized in solution with a 100 fold sequence ($1 \times 10^6$ mass) excess of total poly(A)$^+$ RNA as described in the legend to Figure 2. Corrections were not made for length effects. Assuming that transcription is asymmetric, hybridization values were adjusted for 50% maximum. Values are expressed as equivalent Rot (Britten et al., 1974). [The $Rot_{1/2}$ of an average (40–50%) G + C 1 kb mRNA is calculated to be $\sim 0.5 \times 10^{-3}$ M sec from the data of Bishop et al. (1974) and Axel, Feigelson and Schutz (1976). In Dictyostelium the average mRNA is ~27–28% G + C (Jacobsen, Firtel and Lodish, 1974). Since the rate of DNA-DNA renaturation decreases as the G + C content decreases (see Firtel and Bonner, 1972), we have corrected the $Rot_{1/2}$ for a 1 kb mRNA in Dictyostelium to $\sim 1 \times 10^{-3}$. The absolute $Rot_{1/2}$ should be within the limits of $0.5–1 \times 10^{-3}$.]

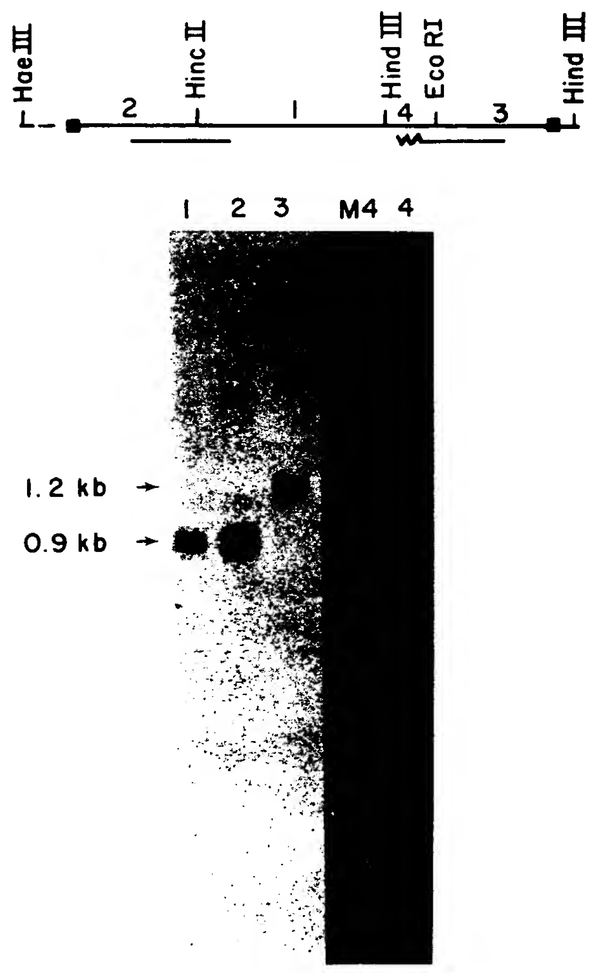


Figure 4. Electrophoresis of $^{32}$P-mRNA Complementary to M4 Clones on Polyacrylamide Gels

DNA was purified from total M4 and M4 subclones and immobilized on nitrocellulose filters. Filters were hybridized with in vivo labeled $^{32}$P-poly(A)$^+$ RNA as described in Table 1. Filters were further washed in 0.05 M NaCl at 45°C. RNA was eluted in 70% formamide, 5 mM EDTA at 55°C and ethanol-precipitated in the presence of carrier tRNA (see Kindle and Firtel, 1978). RNA was size-fractionated on urea-SDS-polyacrylamide gels (Spradling et al., 1977) and RNA bands were identified by autoradiography. Lengths of mRNA were determined by comparison with the mobilities of $^{32}$P-poly(A)$^-$ RNA electrophoresed in adjacent slots. The M4 restriction map and regions transcribed into mRNA are indicated.

When mRNA complementary to band 1 is eluted, it hybridizes to band 2; band 2-complementary RNA cross hybridizes to band 1. Neither mRNA fraction cross-hybridizes to either band 3 or 4. In addition, the results shown in Table 1 indicate that bands 1 and 2 hybridize approximately the same fraction of total $^{32}P$-poly(A)$^+$ RNA (0.1%), and that both hybridizations are partially sensitive to RNAase. From these data we conclude that restriction fragments 1 and 2 comprise a single transcription unit which encodes the 0.9 kb mRNA. It can also be seen in Figure 4 that band 3 hybridizes a 1.2 kb mRNA. On the basis of the fraction of poly(A)$^+$ mRNA present in the eluted RNA, we estimate that the 1.2 kb mRNA is a low abundance class message representing ~0.01% of poly(A)$^+$ mRNA, in agreement with the data presented in Table 1 and Figure 3.

In contrast to the results obtained with subcloned bands 1–3, band 4 complementary mRNA (~1% of total) migrates heterogeneously. Since fragment 4 contains a repeat sequence and since approximately 80–90% of the RNA hybridizing to band 4 is sensitive to low levels of ribonuclease (see Table 1), we suggest that this RNA population consists of a large number of different mRNA molecules containing a short sequence complementary to the repeat found in band 4. The remaining sequences of these mRNA molecules are not complementary to M4 plasmid DNA. They are present as single-stranded tails in the hybrid and are thus sensitive to nuclease digestion. To examine whether these additional RNA sequences are transcribed from single-copy or repeated sequences, M4 band 4 complementary RNA was hybridized to a vast excess of Dictyostelium nuclear DNA. Figure 5 shows that approximately 90% of the mRNA that hybridizes to band 4 is complementary to single-copy DNA ($Cot_{1/2}$ ~120; Firtel and Bonner, 1972). These data therefore suggest that the bulk of the transcribed members of the M4 repeat family are linked to individual single-copy sequences which are also transcribed. The data described below present evidence that the mRNA complementary to band 3 is one of these transcripts, and that the repeat sequence is localized at the 5′ end of this mRNA.

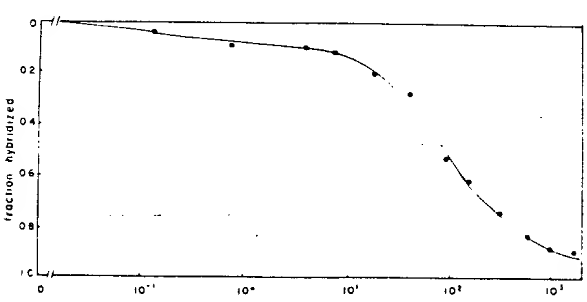


**Figure 5. Hybridization Kinetics of Band 4 Complementary $^{32}P$-Poly(A)$^+$ RNA to a Vast Excess of Dictyostelium DNA**

$^{32}P$-poly(A)$^+$ RNA complementary to band 4 was purified as described in the legend to Figure 4. The RNA was additionally purified by a second cycle of hybridization to band 4 filters. The RNA was hybridized with an excess of genomic DNA and reactions were assayed by nuclease digestion as described previously (Firtel et al., 1972). Values are expressed as equivalent Cot (Britten et al., 1974).

## Repeat Is Asymmetrically Represented in Poly(A)$^+$ RNA

To elucidate the structure and localization of potential transcription units containing the band 4 repeat sequence, it was necessary to determine whether one or both strands of the repeat were transcribed into mRNA. Subcloned band 4 was excised with Eco RI and Hind III, and the ends were labeled using the Klenow fragment of DNA polymerase I and $^{32}P$-dXTPs. Strands were separated on polyacrylamide gels and shown to be ~2% contaminated with their complementary strand (Table 2). Each strand was then hybridized with a $3 \times 10^5$ mass excess of poly(A)$^+$ RNA. The upper band hybridized at a $Rot_{1/2}$ of 0.15 (Figure 6). If we assume the RNA complementary region to be ~0.2 kb long, then ~0.13% of total poly(A)$^+$ RNA is driving the reaction. Furthermore, since the length of the mRNA hybridizing to band 4 is 5–10 times greater than the complementary region (Table 1; Figures 4 and 5), we conclude that 0.5–1.5% of vegetative mRNA molecules contain the repeat sequence. These results agree well with those already presented (see Table 1). We detected no specific hybridization to the lower strand. The upper band complementary transcripts would compete for hybridization to any transcripts which might be complementary to the lower band. The limit of resolution in our experiments is approximately 2% of the sequence concentration of the RNA hybridizing to the upper fragment. If the lower band is transcribed, such transcripts must be present at less than 0.002% of

Table 2. Characteristics of Strand-Separated Band 4 DNA

| | % Hybridization[a] | |
|---|---|---|
| | Upper Band | Lower Band |
| Reaction with subclone 4 DNA[b] | 87 | 91 |
| Zero time binding[c] | 4 | 3 |
| Self-reaction[d] | 4 | 5 |
| Poly(A)$^+$ RNA reaction[e] | | |
| Rot 3 | 75 | 5 |
| Rot 1000 | 80 | 7 |

[a] Percentage of $^{32}P$ strand-separated DNA which binds to hydroxyapatite.
[b] Hybridization with 1 μg of subclone 4 DNA.
[c] Percentage of $^{32}P$ tracer DNA which binds HAP after denaturation by boiling, followed by immediate quenching.
[d] Hybridization in the presence of only carrier tRNA.
[e] Hybridization with $3 \times 10^5$ fold mass excess of poly(A)$^+$ RNA at indicated Rot. Values are not corrected for tracer reactivity or control hybridization.

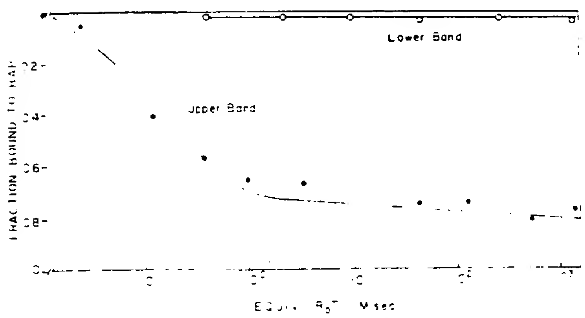


Figure 6. Hybridization of Strand-Separated Band 4 DNAs to Excess Poly(A)$^+$ RNA

Band 4 DNA was end-labeled with $^{32}P$ using the Klenow fragment of DNA polymerase I and strand-separated on polyacrylamide gels. Purified upper and lower bands (spec. act. ~4 × $10^6$ cpm/μg) were each hybridized with $10^5$ fold mass excess of poly(A)$^+$ RNA. Hybridizations were assayed by hydroxyapatite chromatography (Britten and Kohne, 1968). A 0.5 rate correction was made to account for the length difference between the longer driver RNA and shorter tracer DNA (Chamberlin et al., 1978). Hybridization values were corrected for control reactions but not for tracer reactivity. Details are given in Experimental Procedures.

the total poly(A)$^+$ RNA or fewer than ten molecules per cell.

### Bands 3 and 4 Are Part of a Single Transcription Unit

The data presented in the previous sections suggest that a large fraction of the different members of the M4 repeat family and adjacent single-copy sequences are transcribed into mRNA. To determine whether the RNA transcripts complementary to bands 1–3 represent one of these single-copy transcripts, sandwich hybridization experiments were carried out using the method of Dunn and Hassell (1977). DNA of subclones 1, 2 and 3 was digested with Pst I to excise the Dictyostelium insert. The digests were then electrophoresed on agarose gels and transferred to nitrocellulose filters by the method of Southern (1975). The DNA blot filter was then hybridized with unlabeled poly(A)$^+$ RNA (see Figure 7A). If an RNA complementary to band 1–3 contained the repeat sequence found in band 4, it should also be a site for hybridization of a band 4 probe. It can be seen in Figure 7 that the RNA which hybridizes to band 3 will hybridize to nick-translated band 4. Some nonspecific binding of $^{32}P$ band 4 to the pBR322 vehicle is observed, as is binding to bands 1 and 2. This background, however, is considerably less than the specific hybridization of band 4, which is dependent upon the presence of the low abundance class (0.01% of total mRNA) band 3 mRNA. We conclude that band 4 is not part of the 0.9 mRNA. We conclude that band 4 is not part of the 0.9 kb mRNA transcription unit and that the 0.9 kb mRNA complementary to both fragments 1 and 2 does not contain this repeat sequence.

Control experiments in which band 3 RNA hybrids were treated with a low level of ribonuclease prior to the addition of band 4 show that the hybridization of the band 4 probe requires an RNA tail sensitive to RNAase treatment (data not shown). When similar sandwich hybridizations are performed using separated strands of band 4, only the upper strand hybridizes to band 3 complementary RNA (see Figure 7). This result is consistent with the data described previously, indicating the asymmetric transcription of the repeat, and further defines the specificity of the sandwich hybridization. From these data we conclude that fragments 4 and 3 comprise a single transcription unit.

### Repeated Sequence Is Transcribed onto the 5′ End of the RNA

Since band 1 does not hybridize to an RNA which is also complementary to band 4, it is probable that band 4 defines either the 3′ or 5′ end of the band 4–3 mRNA (Figure 7). We have localized the repeat to the 5′ end using strand-specific exonucleases. Subclone 4 was linearized with either Eco RI or Hind III and then digested for various times with either a 3′ strand-specific exonuclease (Exo III) or a 5′ strand-specific exonuclease (λ exonuclease). The extent of digestion was monitored by assaying the liberation of acid-soluble nucleotides from a $^3H$ in vivo labeled pMB9 linear marker included in each digestion. The data are shown in Figure 8. As expected, digestions plateau at ~50% acid solubility since both exonucleases are double-strand-specific. DNA from each point was immobilized on nitrocellulose filters and hybridized with in vivo labeled $^{32}P$–mRNA. As previously shown, band 4 DNA hybridizes to ~1% of total mRNA. Hybridization of $^{32}P$–mRNA to Eco RI-linearized subclone 4 DNA (B) was unaffected by Exo III digestion. In contrast, $^{32}P$–mRNA hybridization to subclone 4, which was linearized with Hind III and digested with Exo III (A), was reduced by 10–20 fold. As expected, when λ exonuclease was used, the complementary hybridization patterns were obtained. These results further demonstrate the asymmetry of transcription of the repeat sequence.

Since the 3′-specific (Exo III) digestion of band 4 from the Eco RI site does not digest the mRNA complementary strand, the Hind III site must lie to the 3′ side of the Eco RI site on the M4 coding strand. Since transcription is antiparallel, the repeat must be located at the 5′ end of the band 4–3 mRNA of plasmid M4.

## Discussion

Short period interspersion, the linkage of short repeat sequences and single-copy DNA, is a feature of most eucaryotic genomes (Davidson et al., 1975a), although there are exceptions (Manning, Schmid and Davidson, 1975; Crain, Davidson and Britten, 1976a; Crain et al., 1976b; Hudspeth, Timberlake and Gold-

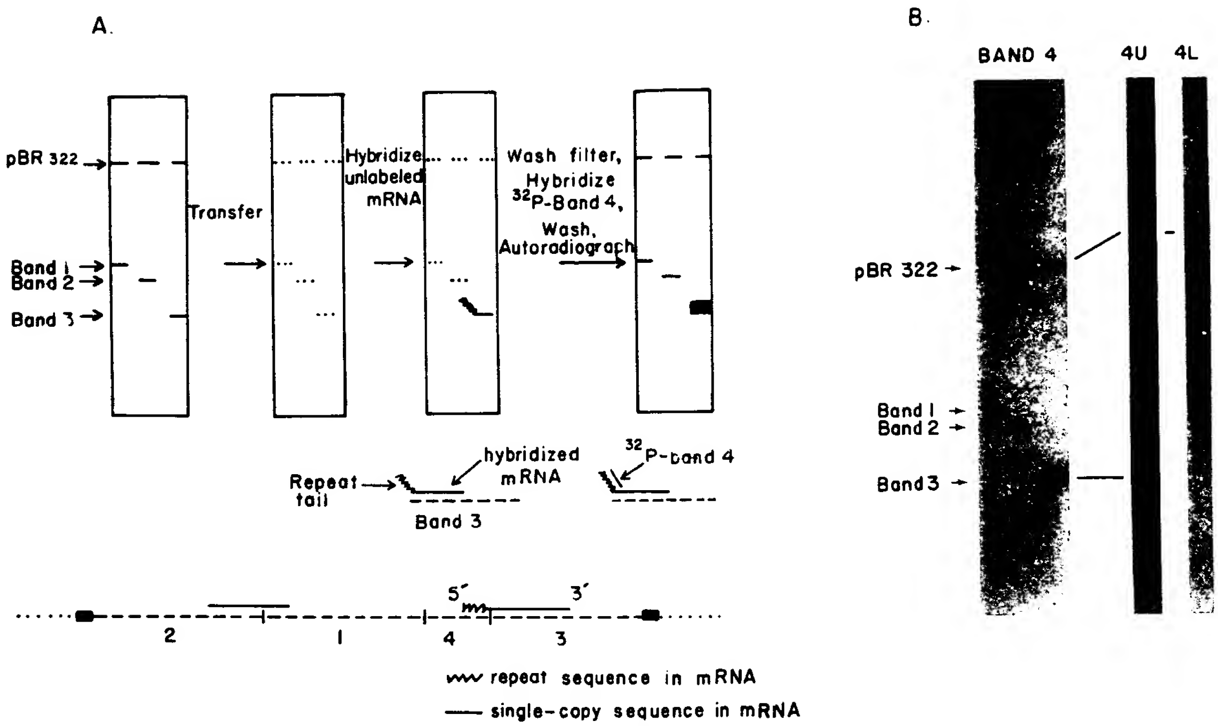


Figure 7. Sandwich Hybridization of $^{32}$P Band 4 Probes to Bands 1, 2 and 3 RNA-DNA Hybrids

(A) Diagram of the sandwich hybridization technique (after Dunn and Hassell, 1977). Bands 1, 2 and 3 were excised from their respective DNA subclones with Pst I and electrophoresed on agarose gels. The DNA was transferred to nitrocellulose filters by the Southern technique and hybridized with unlabeled poly(A)$^+$ RNA. The filters were washed and hybridized with $^{32}$P band 4. Hybridizations were identified by autoradiography. The mRNA hybridizing to band 3 possesses an unhybridized tail which is a site for $^{32}$P band 4 annealing. The restriction map of M4 is presented along with the regions transcribed into mRNA.

(B) Band sandwich hybridization. Sandwich hybridization of $^{32}$P band 4 to Pst I-digested subclones 1, 2 and 3 DNA. Upper (4U) or lower (4L) strand-separated $^{32}$P band 4 sandwich hybridization to Pst I-digested subclone 3 DNA.

bert, 1977). In Dictyostelium approximately 50% of the single-copy DNA is adjacent to short sequences which are repeated in the genome (Firtel and Kindle, 1975). A member of one of these short repeat families is located on the M4 insert. This sequence (~300 bp) is repeated ~100 times in the genome, and in M4 it is interspersed with single-copy DNA. Furthermore, restriction digest patterns of Dictyostelium DNA hybridized to the repeat indicate that the M4 repeat is located on more than 50 restriction fragments ranging in length from ~1 to ~15 kb (see Figure 1). These restriction patterns indicate that individuals of the M4 repeat family are not tightly clustered, but rather are dispersed in the genome and probably associated with single-copy DNA. We have also shown that ~1% of vegetative polysomal poly(A)$^+$ mRNA molecules contain sequences complementary to the M4 repeat. The repeat sequences repr s nt ~10% of each mRNA; th remaining 90% is transcribed from single-copy DNA, further indicating that the M4 repeat family s quences are linked to single-copy DNA. Similar results ar obtained using whole cell poly(A)$^+$ RNA.

V getativ mRNA complementary to the M4 repeat migrates as a mix d size population when electrophoresed on polyacrylamide gels. No discrete bands are observed above this heterogeneous background (see Figure 4). Since the 0.9 kb mRNA in the total M4 mRNA population is observed easily, if an mRNA comprising 0.1% of total mRNA were part of the band 4 mRNA population, it would also have been detected. We suggest that none of the M4 repeat family transcripts are present at a high level; the transcripts of each member of the M4 repeat family are probably of low abundance (~0.01% of total mRNA, five molecules per cell), as observed with the band 4–3 mRNA. The band 4–3 1.2 kb mRNA represents only 1% of the total mRNA which hybridizes to band 4 and thus would not be detected when band 4 mRNA is analyzed by polyacrylamide gel electrophoresis (see Figure 4, slot 4). Since ~1% of the total mRNA hybridizes to the M4 repeat, we suggest that 50–100 different low abundance mRNAs would contain the M4 repeat. Since the M4 repeat is present in 100 copies in the Dictyostelium genome, we propose that a large proportion of the members of this repeat family are transcribed along with their associated single-copy DNA. We have also demonstrated a considerable asymmetric strand representation of the M4 repeat in both

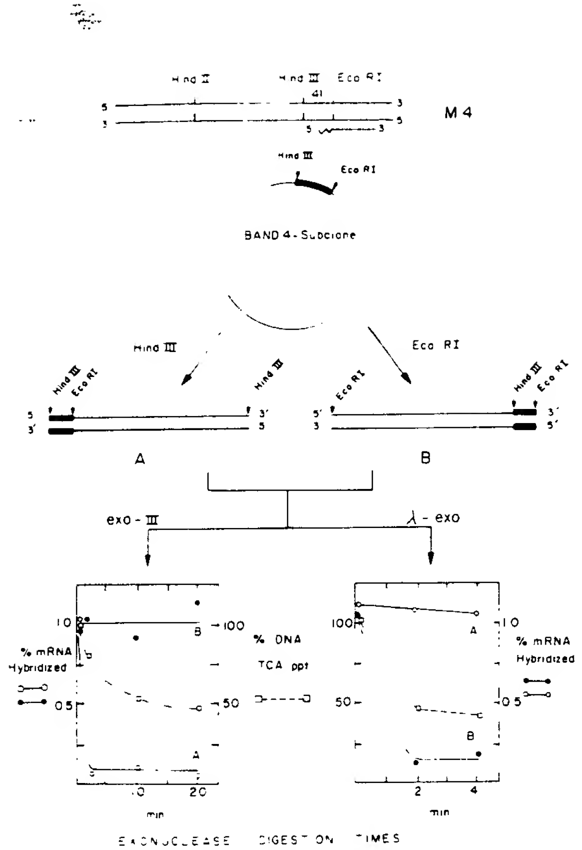


Figure 8. Polarity of the Band 4–3 Transcription Unit

The restriction map of M4 is presented indicating the polarity of the band 4–3 mRNA. The circular map of band 4 subcloned into the Eco RI-Hind III site of pBR322 is also diagrammed. Subclone 4 DNA was linearized with either Hind III (A; ○—○) or Eco RI (B; ●—●). DNA from each reaction was purified and digested for various times with either exonuclease III or λ-exonuclease. Each reaction included in vivo labeled $^3$H-pMB9 linear marker (□—□). Each time point was monitored for TCA precipitability of $^3$H-cpm. DNA from each time point was also immobilized on nitrocellulose filters and hybridized with in vivo labeled $^{32}$P-mRNA as described in Table 1. Hybridizations are not corrected for nonspecific binding to control pBR322 filters.

mRNA and nuclear RNA (A. Kimmel, unpublished results). It is not yet possible to conclude that only one of the strands is transcribed. It is clear, however, that if all members of the M4 repeat family are transcribed with equal frequency, no more than two of the ~100 copies of the M4 repeat are represented in sequences complementary to the lower band.

Another recombinant plasmid (Dd38) has been isolated which contains a different transcribed short repeat linked to single-copy DNA in Dictyostelium. This repeat is also asymmetrically represented on poly(A)$^+$ RNA (C. Lai, unpublished results). It is probable that interspersed repeats are asymmetrically transcribed in Dictyostelium. This observation contrasts sharply with the corresponding one in the sea urchin, where both strands of the repeat families studied thus far are present in nuclear RNA at similar levels [that is, within an order of magnitude (Scheller et al., 1978; Costantini et al., 1978)]. In further contrast, short repeat sequences are probably not present on functional polysomal mRNA in sea urchins (Goldberg et al., 1973; Scheller et al., 1978), although they may be present on maternal mRNA in oocytes (Costantini et al., 1978). It is unclear whether many organisms contain interspersed repeat and single-copy sequences on mRNA. None have been detected in HeLa (Klein et al., 1974) or rodent cells (Campo and Bishop, 1974), although they may be present on mRNA in Xenopus embryos (Dina, Crippa and Beccari, 1973; Dina, Meza and Crippa, 1974). The interspersion of repeat and single-copy complementary sequences in nuclear RNA, however, is common to many eucaryotes (for review see Lewin, 1975).

In addition to the transcription of the repeat, two single-copy regions of M4 are also transcribed. Bands 1–2 code for a moderately abundant (0.1% of total mRNA, ~50 molecules per cell) mRNA and band 3 is transcribed into a low abundance class mRNA. Sandwich hybridization has shown that the repeat is part of the band 3 transcript. Since band 1 is not complementary to an RNA which contains the repeat, the band 4–3 transcription unit must terminate in band 4 (see Figure 7). The two transcription units on M4 are separated by ~1.5 kb.

On the basis of exonuclease hybridization studies, we have concluded that the repeat encodes the 5′ end of the band 4–3 mRNA. Since transcription of the repeat is asymmetric, we propose that polarity of transcription is identical in all M4 repeats transcribed and that the repeat is located at the 5′ end of the set of single-copy transcripts. We propose that transcription initiates in these repeat sequences and elongation proceeds into the single-copy region. Sequences complementary to the repeat are located at the 5′ end of the resulting transcripts, as shown in Figure 8. We do not yet know whether the primary nuclear transcripts contain sequences which are removed and not present in the mature mRNAs. The location of the repeat sequence at the 5′ end of a set of single-copy genes suggests that the repeat sequence may be a regulatory recognition site similar to that postulated by Britten and Davidson (1969) (see also Davidson and Britten, 1973; Davidson, Klein and Britten, 1977). In this model, an entire family of repeat sequences and their corresponding set of single-copy DNA may be similarly regulated during development. Although no direct evidence to support this model is presently available, several interesting correlations have been observed in the sea urchin system. The Davidson and Britten laboratories have observed that most genes active during gastrula are single-copy and adjacent to short repeat sequences (Davidson et al., 1975b). In addition, they have observed that the short repeats are differentially expressed in various sea urchin tissues and developmental stages (Scheller et al., 1978). While we suggest a role for the repeat in the transcriptional regulation of gene activity, we cannot exclude

the possibility that the repeat may be involved in processing, transport or translation of the mRNAs.

It should also be emphasized that the M4 band 1-2 mRNA is encoded entirely by single-copy DNA. Thus not all transcription units in Dictyostelium are composed of interspersed repeat and single-copy sequences. Whatever the function of repeat sequences, they are not essential for expression of all Dictyostelium genes.

To resolve some of the speculations concerning the function of the repeat, additional clones containing different members of repeat families must be characterized. We are presently initiating this work to examine the transcription of single-copy regions associated with the different members of the M4 repeat family and to study the sequences of the repeats to identify possible initiation sites of transcription.

## Experimental Procedures

### Strains and Culture Conditions

Dictyostelium discoideum axenic mutant AX-3 derived from wild-type strain NC-4 (Raper, 1935) by Loomis (1975) was used exclusively in these studies. Cells were grown axenically in Mes-HL-5 (Firtel and Lodish, 1973). For labeling, exponential cells (3-4 $\times 10^6$ per ml) were washed and resuspended in phosphate-free PDF containing MES (Firtel, Jacobson and Lodish, 1972) at $1 \times 10^7$ cells per ml and incubated for 1 hr. The cells were concentrated to $2 \times 10^7$ cells per ml in MES/PDF and labeled for 2-3 hr with 1-2 mCi $^{32}P-O_4$ per ml. Under these labeling conditions, >90% of the total $^{32}$P-poly(A)$^+$ RNA is located on polysomes (R. Firtel, unpublished data).

Plasmids were maintained in E. coli strains C600 ($r_k^-$, $m_k^-$), HB101 (H. Boyer, personal communication) or SF8 (Cameron et al., 1974). Bacteria were grown in M9 casaminoacid media (Roberts et al., 1963).

### Construction of Recombinant Plasmids

M4 contains a 4.5 kb fragment of randomly sheared Dictyostelium nuclear DNA inserted into the Eco RI site of pMB9 by poly(dA-dT) homopolymer extension (K. Kindle and R. Firtel, manuscript submitted).

Restriction fragments 1, 2 and 3 of M4 (see Figure 1) were purified and poly(dC)-extended using terminal transferase (Roychoudhury, Jay and Wu, 1976) as described by W. Rowekamp and R. Firtel (manuscript in preparation). The DNAs were mixed with Pst I-cleaved pBR322 whch had been tailed with poly(dG) (A. Otsuka, manuscript submitted; W. Rowekamp and R. Firtel, manuscript in preparation) and used to transform E. coli. Clones were selected which were Tet$^R$ and Amp$^S$.

To subclone band 4, M4 and pBR322 were mixed at 20:1 molar ratio, digested with Eco RI and Hind III and ligated with T4 DNA ligase. Transformants were selected which were Amp$^R$ and Tet$^S$.

### Preparation of DNA

Dictyostelium genomic DNA was extracted from isolated nuclei as described (Firtel et al., 1976b).

Plasmid DNA was isolated from stationary phase cells using CsCl-ethidium bromide equilibrium centrifugation of a lysozyme-treated, sarkosyl total cell lysate.

DNA fragments were prepared from agarose gels by a modification of the freeze-squeeze method (Cockburn et al., 1976a) and from polyacrylamide gels by elution of gel slice homogenates in 0.4 M LiCl, 0.01 M Tris, 0.001 M EDTA (pH 7.4).

### Preparation of RNA

RNA was isolated from whole cells and the poly(A)$^+$ fraction was purified by passage over poly(U)-Sepharose (Firtel and Lodish, 1973).

$^{32}$P-RNA was extracted from polysomes and $^{32}$P-mRNA was purified by poly(U)-Sepharose affinity chromatography.

$^{32}$P-mRNA was eluted from DNA filters which were washed in 0.05 M NaCl at 45°C by several incubations in 70% formamide, 5 mM EDTA (pH 7.5) at 55°C as described by Kindle and Firtel (1978).

### In Vitro Labeling

DNA fagments were purified and labeled to high specific activity (~2-8 $\times 10^7$ cpm/$\mu$g) by nick translation (Schachat and Hogness, 1973; Maniatas et al., 1975) with DNA polymerase I (Boehringer-Mannheim) and $\alpha$-$^{32}$P-deoxyribonucleotide triphosphates.

In some cases, DNA restriction fragments were end-labeled using DNA polymerase I (Klenow) and $\alpha$-$^{32}$P-dXTPs. Specific activities of ~4 $\times 10^6$ were reproducibly obtained.

### Gel Electrophoresis of DNA and RNA

Purified $^{32}$P-mRNAs were analyzed on urea-SDS-polyacrylamide gels (Spradling et al., 1977).

Restriction enzyme-digested DNAs were electrophoresed on agarose gels and bands were visualized by ethidium bromide staining as described previously (Firtel et al., 1976a).

DNA strands were separated on polyacrylamide after denaturation in 50% DMSO at 90°C for 2 min (Szalay, Grohman and Sinsheimer, 1977).

### Hybridizations

Dictyostelium DNA blot filters (Southern, 1975) were hybridized with nick-translated $^{32}$P-probes as described (Kindle and Firtel, 1978).

Solution hybridizations with nick-translated probes were performed in 270 mM NaCl, 10 mM TES, 1 mM EDTA, 0.2% SDS at 58°C in sealed microcapillary pipettes and assayed by S1 nuclease digestion (Firtel et al., 1972). Self-renaturation of nick-translated probes was monitored in control hybridizations by substituting tRNA for Dictyostelium driver nucleic acid.

RNA excess hybridizations to separated strands of band 4 (see Figure 1) were performed in 0.5 M PB, 1 mM EDTA at 55°C and assayed by hydroxyapatite chromatography as described by Costantini et al. (1978) and Scheller et al. (1978).

DNA blot filters to be used for sandwich hybridizations were pretreated in hybridization buffer [2X Denhardt solution (Denhardt, 1966), 50% formamide, 4 x SSC, 0.12 PB, 10 mM EDTA, 0.2% SDS] at 37°C. The filters were prehybridized with poly(A)$^+$ RNA (0.5 mg/ml) and poly(A) (1 mg/ml) at 37°C for 36 hr. Filters were washed extensively at criterion and incubated with $^{32}$P probes in hybridization buffer at 37°C. Filters were washed again at criterion and hybridization was assayed by autoradiography.

Immobilized DNA filters were prepared as described (Kindle and Firtel, 1978) and hybridized with $^{32}$P-mRNA in sandwich hybridization buffer (see Table 1).

### Acknowledgments

We are indebted to Drs. Karen Kindle and Walter Rowekamp for their help in the initial phase of this work. We also thank Stephen Poole, Andrew Cockburn, Ann McNelly and Cary Lai for their comments.

A.R.K. is a postdoctoral fellow of the American Cancer Society. R.A.F. is the recipient of an American Cancer Society Faculty Research Award. This work was supported by additional grants from the NIH and the NSF.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received December 26, 1978; revised January 22, 1979

### References

Axel, R., Feigelson, P. and Schutz, G. (1976). Cell *7*, 247-254.

Bishop, J. O., Morton, J. G., Rosbash, M. and Richardson, M. (1974). Nature *250*, 199-204.

Britten, R. J. and Kohne, D. E. (1968). Science *161*, 529–540.

Britten, R. J. and Davidson, E. H. (1969). Science *163*, 349–357.

Britten, R. J., Graham, E. E. and Neufeld, B. R. (1974). In Methods in Enzymology, *29E*, L. Grossman and K. Moldave, eds. (New York: Academic Press), pp. 363–418.

Cameron, J. R., Panasenko, S. M., Lehman, I. R. and Davis, R. W. (1974). Proc. Nat. Acad. Sci. USA *72*, 3416–3420.

Campo, M. S. and Bishop, J. O. (1974). J. Mol. Biol. *90*, 649–663.

Chamberlin, M. E., Galau, G. A., Britten, R. J. and Davidson, E. H. (1978). Nucl. Acids Res. *5*, 2073–2094.

Cockburn, A. F., Newkirk, M. J. and Firtel, R. A. (1976a). Cell *9*, 605–613.

Cockburn, A., Frankel, G., Firtel, R. A., Kindle, K. and Newkirk, M. J. B. (1976b). In ICN-UCLA Symposia on Molecular and Cellular Biology, *5*, D. P. Nierlich, W. J. Rutter and C. F. Fox, eds. (New York: Academic Press), pp. 599–603.

Cockburn, A., Taylor, W. and Firtel, R. A. (1978). Chromosoma *70*, 19–29.

Costantini, F. D., Scheller, R. H., Britten, R. J. and Davidson, E. H. (1978). Cell *15*, 173–187.

Crain, W. R., Davidson, E. H. and Britten, R. J. (1976a). Chromosoma *59*, 1–12.

Crain, W. R., Eden, F. C., Pearson, W. R., Davidson, E. H. and Britten, R. J. (1976b). Chromosoma *56*, 309–326.

Davidson, E. H. and Britten, R. J. (1973). Quart. Rev. Biol. *48*, 565–613.

Davidson, E. H., Klein, W. H. and Britten, R. J. (1977). Dev. Biol. *55*, 69–84.

Davidson, E. H., Galau, G. A., Angerer, R. C. and Britten, R. J. (1975a). Chromosoma *51*, 253–259.

Davidson, E. H., Hough, B. R., Klein, W. H. and Britten, R. J. (1975b). Cell *4*, 217–238.

Denhardt, D. (1966). Biochem. Biophys. Res. Commun. *23*, 640–646.

Dina, D., Crippa, M. and Beccari, E. (1973). Nature New Biol. *242*, 101–105.

Dina, D., Meza, I. and Crippa, M. (1974). Nature *248*, 486–490.

Dunn, A. R. and Hassell, J. A. (1977). Cell *12*, 23–36.

Firtel, R. A. and Bonner, J. (1972). J. Mol. Biol. *66*, 339–351.

Firtel, R. A. and Lodish, H. F. (1973). J. Mol. Biol. *79*, 295–314.

Firtel, R. A. and Kindle, K. (1975). Cell *5*, 401–411.

Firtel, R. A., Jacobson, A. and Lodish, H. F. (1972). Nature New Biol. *234*, 225–228.

Firtel, R. A., Cockburn, A., Frankel, G. and Hershfield, V. (1976a). J. Mol. Biol. *102*, 831–852.

Firtel, R. A., Kindle, K. and Huxley, M. P. (1976b). Fed. Proc. *35*, 13–22.

Goldberg, R. B., Galau, G. A., Britten, R. J. and Davidson, E. H. (1973). Proc. Nat. Acad. Sci. USA *70*, 3516–3520.

Grunstein, M. and Hogness, D. S. (1975). Proc. Nat. Acad. Sci. USA *72*, 3961–3965.

Hudspeth, M. E. S., Timberlake, W. E. and Goldberg, R. B. (1977). Proc. Nat. Acad. Sci. USA *74*, 4332–4336.

Jacobsen, A., Firtel, R. A. and Lodish, H. F. (1974). J. Mol. Biol. *82*, 213–230.

Kindle, K. L. and Firtel, R. A. (1978). Cell *15*, 763–778.

Klein, W. H., Murphy, W., Attardi, G., Britten, R. J. and Davidson, E. H. (1974). Proc. Nat. Acad. Sci. USA *71*, 785–789.

Lewin, B. (1975). Cell *4*, 77–93.

Loomis, W. F. (1975). *D. discoidium*: A Developmental System (New York: Academic Press).

Maizels, N. (1976). Cell *9*, 431–438.

Maniatis, L., Jeffrey, A. and Kleid, D. G. (1975). Proc. Nat. Acad. Sci. USA *72*, 1184–1188.

Manning, J. E., Schmid, C. W. and Davidson, N. (1975). Cell *4*, 141–155.

Morrow, J. F. (1974). Ph.D. thesis, Stanford University, Stanford, California.

Raper, K. B. (1935). J. Agricul. Res. *50*, 135–141.

Roberts, R. B., Abelson, P. H., Crowie, D. B., Bolton, E. T. and Britten, R. J. (1963). Carnegie Inst. Washington Yearbook *61*, 607.

Roychoudhury, R., Jay, E. and Wu, R. (1976). Nucl. Acids Res. *3*, 101–116.

Schachat, F. H. and Hogness, D. S. (1973). Cold Spring Harbor Symp. Quant. Biol. *38*, 371–381.

Scheller, R. H., Costantini, F. D., Kozlowski, M. R., Britten, R. J. and Davidson, E. H. (1978). Cell *15*, 189–203.

Smith, M. J., Britten, R. J. and Davidson, E. H. (1975). Proc. Nat. Acad. Sci. USA *72*, 4805.

Southern, E. M. (1975). J. Mol. Biol. *98*, 504–517.

Spradling, A., Pardue, M. L. and Penman, S. (1977). J. Mol. Biol. *109*, 559–587.

Szalay, A. A., Grohman, K. and Sinsheimer, R. L. (1977). Nucl. Acids Res. *4*, 1569–1578.